



LUND UNIVERSITY

Kinds of Learning and the Likelihood of Future True Beliefs: Reply to Jäger on Reliabilism and the Value Problem

Olsson, Erik J; Jönsson, Martin

Published in:

Theoria: a Swedish Journal of Philosophy

DOI:

[10.1111/j.1755-2567.2011.01103.x](https://doi.org/10.1111/j.1755-2567.2011.01103.x)

2011

[Link to publication](#)

Citation for published version (APA):

Olsson, E. J., & Jönsson, M. (2011). Kinds of Learning and the Likelihood of Future True Beliefs: Reply to Jäger on Reliabilism and the Value Problem. *Theoria: a Swedish Journal of Philosophy*, 77(3), 214-222.
<https://doi.org/10.1111/j.1755-2567.2011.01103.x>

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Kinds of Learning and the Likelihood of Future True Beliefs: Reply to Jäger on Reliabilism and the Value Problem

Erik J. Olsson and Martin L. Jönsson

Olsson, E. J. & Jönsson, M. (2011). Kinds of Learning and the Likelihood of Future True Beliefs: Reply to Jäger on Reliabilism and the Value Problem. *Theoria*, 77, 214-222. Wiley-Blackwell.

Abstract: We reply to Christoph Jäger’s criticism of the conditional probability (CP) solution to the value problem for reliabilism due to Goldman and Olsson (2009). We argue that while Jäger raises some legitimate concerns about the compatibility of the CP solution with externalist epistemology, his objections do not in the end reduce the plausibility of that solution.

1. Introduction

If a belief is already true, adding that it was reliably produced doesn’t seem to make it more valuable. Hence, reliably acquired true belief is no more valuable than mere true belief. Knowledge, by contrast, *is* more valuable than mere true belief. But if so, (simple) reliabilism must be wrong, for it equates knowledge with reliably acquired true belief. This, in a nutshell, is the value problem for reliabilism. It is sometimes called the swamping problem because, as just indicated, the value of true belief seems to “swamp” the value of reliable acquisition.¹

In Goldman and Olsson (2009), two ways of addressing the value problem from the perspective of reliabilism were proposed. According to the *conditional probability (CP) solution*, reliabilist knowledge is more valuable because it makes future true belief more probable. This value is attained normally and not in every single case. The second proposal, *type-instrumentalism and value autonomization (TIVA)*, was intended to address, among other things the issue why some people tend to think that knowledge is

¹ See, for instance, Kvanvig (2003) for the swamping problem and various approaches to it.

always more valuable than mere true belief. While Olsson is more inclined to defend, the CP solution, Goldman has reported that he favors the TIVA solution.²

In his paper “Reliabilism and the Value Problem”, Christoph Jäger objects to the CP solution, arguing that it is incompatible with epistemological externalism. This, if true, would be unfortunate for that proposal because it would be deemed unacceptable, or at least highly problematic, by the majority of reliabilists who are also externalists.

In this article we take the opportunity of responding to Jäger’s stimulating criticism. Before we do that we will briefly set the stage by stating the CP solution in more precise terms as well as adding some general remarks concerning its scope and limits.

2. The conditional probability solution in outline

The basic idea behind the CP solution is that a state of knowledge is also a state of reliable acquisition and as such it is valuable not only as an indicator of the truth of the belief thus acquired but also as indicative of the production of further true beliefs (of a similar kind), namely, true beliefs resulting from reapplications of the reliable method in question. The basic CP claim, then, is that the probability of future true belief (of a similar kind) is greater conditional on reliabilist knowledge than it is conditional on mere true belief. More formally,

(CP) $P(S \text{ acquires further true beliefs of a similar kind} \mid S \text{ has reliabilist knowledge that } p) > P(S \text{ acquires further true belief of a similar kind} \mid S \text{ has a mere true belief that } p)$.

The (CP) claim being true would be a reason why knowledge is more valuable, indeed *epistemically* more valuable, than mere true belief.³

But what reasons are there to believe that (CP) is actually true? The answer is that it is true in virtue of certain empirical conditions characterizing our world: *non-uniqueness*,

² What follows is our own view on these matters, which should not automatically be taken to correspond to Goldman’s. For a recent statement of Goldman’s position on the swamping problem and the value of knowledge, see Goldman (2009), pp. 254-262.

³ Olsson (2009) features a brief discussion of the compatibility of the CP solution with externalism occasioned by objections raised by Jäger at the 2008 Dusseldorf symposium on Goldman’s work. In the same article, Olsson also considers various other objections that have been raised against the CP solution.

cross-temporal access, learning and generality. By non-uniqueness, the same kind of problem will tend to arise more than once. Once you encounter a problem of a certain type, you are likely to encounter a problem of the same type at some later point. This is true, for instance, of the problem of finding the way from A to B. By cross-temporal access, a method that was used once will tend to be available when the problem arises in the future. The method of using the GPS navigation device in your car would be a case in point. By the learning assumption, if a particular method solves a problem once, and you have no reason to think that it did so unsuccessfully, you will tend to use the same method again. By generality, finally, if a method was reliable in the past, its reliability is unlikely suddenly to be discontinued. Rather, the method is likely to remain reliable. Again, using a GPS device provides a good illustration. Once you get a feel for how to use it, and you have no reason to think it has failed, you will use it again and again, and the device will continue to be reliable – at least up to a point.

It must be acknowledged that there are many exceptions to these regularities, e.g. problems arising only once in a lifetime, malfunctioning GPS devices, and so on. But the fact that there are such exceptions is compatible with (CP). For what (CP) states is only that knowledge *raises the probability* of future true belief in comparison to mere true belief; it doesn't state that knowledge *guarantees* future true belief. It follows that the extra value of knowledge supported by (CP) is only a kind of *pro tanto* or default value: in the absence of any positive reason to believe that the empirical conditions just mentioned were violated in the particular case in question, we may assume, or take for granted, that knowing that p is more valuable than merely believing truly that p.⁴

3. Jäger's initial worry

What has been said so far is sufficient to provide the background against which an “initial worry” of Jäger's can be addressed:

⁴ In Olsson's view, reliabilist knowledge has a stability value that goes beyond the “repetition value” accounted for by the CP solution: the fact that a person knows that p, in the reliabilist sense, as opposed to merely truly believes that p, makes that person's belief that p more stable. More precisely, the probability that a given true belief will remain in place is greater given that it was reliably, as opposed to unreliably, acquired. Stability, as Williamson (2000) notices, is practically valuable in relation to actions that take place over an extended period of time, but it could also be argued that there is a distinct *epistemic* benefit to having true beliefs that are stable. For the details, see Olsson (2007) and (2008).

An initial worry is that nothing in Goldman and Olsson's account rules out that, if some reliable epistemic mechanism R were not operating in a given belief forming process, this might cause a still more reliable mechanism R* to operate, not on the present, but on the relevant future occasions. In that case, it is more probable that S acquires more true belief on those occasions if on the present occasion S does not employ R but acquires – for example by sheer guesswork – a true belief that has not been produced by any reliable mechanism.

Jäger provides the following thought experiment in support of his contention:

Suppose that on your way to Larissa you are using a navigation system N1 which, by suitable reliabilist standards, counts as reliable. Nevertheless, let us suppose that N1 gives you the correct route at the first and second crossroads, but not at the following three. Suppose further that, had you not been using this system at the first crossroads but instead just guessed correctly where to go (for example, because you believed that the system was broken), you would have stopped at the next service station and bought a top quality system N2 that is much more reliable than N1. In that case, the probability that you acquire more true geographical beliefs for the rest of your trip is greater than in a situation in which you use the low quality system N1 at the first crossroads, and continue to use it. So employing the reliable mechanism N1 does not raise but instead lowers the conditional probability that you acquire more true beliefs of a similar kind in the future.

An immediate problem with this example is that the learning condition is not satisfied in the hypothetical scenario in which you simply guess (correctly) where to go, buy the top quality N2 system and uses that system from there on. The switch from guessing to using N2 violates the learning condition because it constitutes an unprovoked shift from one method to another. Hence, what Jäger's example shows, as it stands, is that reliabilist knowledge may fail to make future true belief more likely in special cases in which

learning is not satisfied. But that this is so was acknowledged already in Goldman and Olsson (2009).⁵

However, there is a twist to the story. It may be argued that although Jäger's thought experiment is in violation of the learning condition, the core of his argument does not really depend on that violation. Consider the following alternative thought experiment: You have a faulty GPS device (N0), but, being almost broke, the new one that you can afford (N1) will only be somewhat reliable (but still reliable). However, if you continue using the old one until it breaks down completely your insurance company will replace it with one that is extremely reliable (N2). Assume that you have registered the fact that N0 is unreliable, so that regardless of what you do next you won't violate learning (since learning only requires you to *retain a procedure if you haven't encountered any problems with it*). Now, consider two scenarios. In scenario 1, you buy N1. In scenario 2, you continue to use N0 until it breaks down completely (in which case you receive N2 from the insurance company). As a consequence, there is a time *t* at which you use reliable, but not entirely trustworthy, N1 in scenario 1 but still use unreliable N0 in scenario 2. Suppose that, at *t*, you form a true geographical belief that *p* by means of consulting your GPS system. But if so, the reliabilist knowledge present in scenario 1 fails to make future true belief more likely, for after a while the unreliable method used in scenario 2 will be replaced by super reliable N2, which will outperform N1 used in scenario 1 in the long run.

In other words, reliabilist knowledge may fail to make future true belief more likely, even if the learning condition is satisfied, and even if the other empirical conditions are satisfied as well. This is a novel observation that is worth making. Nonetheless, it doesn't threaten the CP solution to the value problem. The CP claim is a statistical claim about the relation between two conditional probabilities or, in frequentist terms, proportions. As such it is not jeopardized by reference to certain special cases involving some rarely occurring happenings, and this is so even if the empirical conditions that are central to the CP solution are not violated in those cases.

4. Jäger on the CP solution and externalism

⁵ Ibid., p. 30.

Let us now turn to Jäger's main critique. Jäger's point of departure is the following quotation from Goldman and Olsson (2009) explaining the learning condition in the context of a modernized version of Plato's Larissa example (from the dialogue *Meno*):

if you have used a given method before and the result has been unobjectionable, you are likely to use it again on a similar occasion, if it is available. Having invoked the navigation system once without apparent problems, you have reason to believe that it should work again. Hence, you decide to rely on it also at the second crossroads (p. 13).

Jäger now observes that this illustration of the learning condition is couched in terms that look internalistically loaded. Indeed, Goldman and Olsson seem to be saying, or implying, that an organism can satisfy the learning condition only if it is capable of having *reasons to believe* that the method in question should work again, something which betrays the externalist spirit of reliabilism. Based on considerations such as these, the details of which need not concern us, Jäger proceeds to conclude that the learning condition presupposes that the subject in fact has various beliefs about the method used and its reliability:

CPS, Goldman and Olsson concede, works only if the Learning condition is satisfied, but Learning presupposes that S believes (i) that the mechanism employed in a given belief forming situation delivered a true belief, and (ii) that that process or method was, and still is, reliable.

But if this is true then reliabilist knowledge has an extra value according to the CP solution only if some internalist conditions obtain. In Jäger's own words,

[h]ence on Goldman and Olsson's account acquiring reliabilist knowledge on a given occasion has more value for S than merely acquiring a corresponding true belief, only

if S believes that, on that occasion, S has acquired a true belief that was produced by a reliable mechanism.

The upshot is that the CP solution thus seems to be incompatible with the externalistic spirit of reliabilism.

Jäger bases this interpretation of the learning condition on the Larissa example, which was used as a paradigm case in Goldman and Olsson's text. The Larissa example involves the higher level task of using a GPS device for finding the way to Larissa. In that particular context, it is natural to couch the discussion in terms of "decisions", "reasons to believe" et cetera. Jäger is correct in noting that those terms are suspect from an externalist point of view. In other words, extrapolating a general theory about these matters from a literal interpretation of what Goldman and Olsson wrote in connection with the Larissa example is problematic. That much seems true.

As Jäger acknowledges, the more fundamental issue, however, is to what extent the empirical assumptions underlying the CP solution can be cast in a way that secures the extra value of knowledge without presupposing higher-level cognition on the part of the believer. Jäger offers the following proposal in this regard:

One move that springs to mind is to replace the current version of the Learning Condition with the claim that people are simply hardwired to reuse epistemic processes that have proven to be epistemically 'unobjectionable'. Such a wired-in disposition would still constitute some kind of second-order epistemic mechanism.... However, a mechanism M of the desired kind would have to be able to distinguish reliable epistemic processes from unreliable one. Why else should it foster re-employment of the former rather than the latter? ... In order for M to identify a method as reliable, M would apparently have to rely, not on a single success, but on a favourable track record of employments of that process or method. M would have to contain some kind of unconscious inductive mechanism that would generate a tendency or disposition in the subject to reuse epistemic processes which over a representative series of trials it records as successful, while at the same time blocking tendencies to reuse processes with an unfavorable track record.

Jäger concludes that “the prospects for developing a convincing proposal along such lines do not appear rosy”. His main reason for believing this, it seems, is that he does not know of any evidence supporting that a mechanism such as M exists.

Jäger’s argument makes good sense on the following understanding of the learning condition:

(Strong Learning)

- a) People have a disposition to reuse a belief formation process if that process has been shown to be reliable.
- b) People have a disposition to replace a belief formation process if that process has been shown to be unreliable.

For if strong learning were to be compatible with externalism then there would have to exist some kind of unconscious inductive mechanism that i) would generate a tendency or disposition in the subject to reuse epistemic processes which over a representative series of trials it records as successful” (due to a)) and ii) would block tendencies to reuse processes with an unfavorable track record (due to b)).

However, the problem with Jäger’s argument is that Strong Learning is a much stronger assumption than what is needed for the purposes of the CP solution. What is needed is only a learning condition of the following kind:

(Learning) People have a disposition to reuse a belief formation process if that process has not been shown to be unsuccessful.

Learning is removed from strong learning on three counts; i) the a) clause of strong learning presupposes that a certain condition must be satisfied in order for an agent to carry on using a process, while learning entails that an agent will carry on using a process at least until some condition obtains, ii) according to strong learning retention is conditional on *reliability*, but according to learning retention lasts at least until

unsuccessfulness obtains, and iii) there is nothing in learning corresponding to the b) clause of strong learning (i.e. no thesis concerning when an agent will change belief forming process). Let's go through these points in order.

First, in contrast with the a) clause of strong learning, learning does not require that the reliability of a process be demonstrated prior to the reemployment of that process. In fact, it doesn't require that anything be demonstrated before a process can be reused. In order to satisfy learning, it is sufficient that an organism is hard-wired in accordance with the law of least resistance, the most economical way of reacting to stimuli being to continue responding in the old way rather than trying something new. This clearly does not require higher order powers of representation and it is clear that not only humans, but also many cognitively considerably simpler organisms can be hard-wired in this way.⁶

Stubbornly repeating the same pattern of action is not always the best approach to life's challenges, which is why it seems reasonable to assume also that many creatures are equipped with sensors and other mechanisms needed for signaling to the brain when the organism's desires have been frustrated. Once those mechanisms fire, the organism would typically react differently the next time around, although exactly how it will respond to future stimuli is not explicitly specified in the learning condition. It is important to note how little is required in order for CP to work. For instance, if all agents retained their reliable belief forming processes no matter what this would *strengthen* the connection between having a reliably formed belief and potentially having more true beliefs in the future. As long as the other empirical conditions obtain, die-hard retention would, in normal circumstances, lead to an increased probability of future true beliefs. What CP needs, and what learning provides, is only the requirement that agents often retain their belief forming processes.

⁶ Jäger briefly mentions a value ascription problem that – if we understand him correctly – is supposed to follow from the above account in terms of hard-wired dispositions to retain belief forming procedures. Jäger points out that the object of value is 'an arbitrary, single instance of...reliabilist knowledge' and not a series of true beliefs produced by the same reliable process. But – again, if we understand him correctly – if one adopts a hardwired proposal along the lines of strong learning then it might seem that this is entailed since strong learning requires a process to generate a series of true beliefs before the continued employment of the process is secured. So arbitrary reliably true beliefs (such as the first ones produced by a process) do not have an extra value over merely true beliefs since the retention of a belief forming process is not guaranteed by learning until we have a sufficiently strong track record. However, note that this objection only follows from strong learning and not from learning. If learning only entails something like the law of least resistance, it does provide support for an arbitrary reliably produced true belief, since retention here is not conditional on there being any track record at all.

Second, it is not required that the condition that is invoked in learning is “detected reliability” or “detected unreliability”. If humans always retained their belief forming processes regardless of whether they were reliable or unreliable this would be unproblematic from the perspective of CP. So exactly under which conditions it is unreasonable to expect that agents will retain their belief forming procedures is not essential to CP. So neither detecting whether a process is reliable or unreliable is required from the perspective of CP. So something much weaker like ‘unsuccessful’ is sufficient. This might be a little vague, but this is appropriate since i) it doesn’t matter from the perspective of CP *exactly* how it is spelled out (as long as it doesn’t entail that agents abandon their belief-forming processes helter skelter), and ii) exactly what retention is actually conditional on is not a philosophical question but an empirical one. Again, what CP needs, and what learning provides, is only the requirement that agents often retain their belief forming processes.

Third, as have already been emphasized, there is nothing in learning that corresponds to the b) clause of strong learning. In other words, there is nothing at all in the learning condition that entail that agents ever change belief forming processes. So Jäger is mistaken in thinking that it follows from CP that organisms are required to have “some kind of unconscious inductive mechanism that would generate a tendency or disposition [to block] tendencies to reuse processes with an unfavorable track record.”

However, it might be worth mentioning that even if CP did contain such a commitment this might not be unreasonable in light of available developmental data.

For instance, during language learning children makes use of a process such that whenever they encounter an unfamiliar word they will assume that it stands for something distinct from what the other words that they possess stands for (Golinkoff et al. 1994). This heuristic (sometimes called the *novel name – nameless category principle*) is a process that will initially lead to mostly true beliefs about word meanings, but eventually, the speaker, having encountered numerous counter examples, will abandon the process. The abandonment of this process is not a conscious decision and most adults do not know that they ever made use of the process.

It is hard to say with certainty that this, or other shifts during development, are due to sensitivity to unfavorable track records or whether they simply reflect the cognitive

maturation of the individual. However, there are many examples of these kinds of transitions, and a *prima facie* reasonable explanation of a shift in a rational agent from an unreliable to a reliable process involves the unreliability of the former. If we explain the shift without reference to unreliability, the accidental cognitive improvement that results from the shift needs separate explanation. So, to sum up, even if learning did presuppose the mechanism M, this does not seem very problematic since its existence seems *prima facie* reasonable in the light of developmental studies. However, as we have emphasized the existence of M only follows from something like strong learning.

To sum up these three points, what is needed for the CP solution to work is much less than strong learning. What is needed is only a weaker learning condition that is, in all probability, externalist friendly.

5. Conclusion

Cristoph Jäger should be credited for questioning the compatibility of the CP solution with externalist epistemology. In particular, he is right to point out that Goldman and Olsson's original formulation of the empirical conditions needed for the CP solution to take off suggested that the satisfaction of those conditions would require cognitive access to methods and their reliability. However, in this reply we have shown that those conditions can in fact be cast in an externalist-friendly manner. Our conclusion is that the CP solution to the value problem is compatible with externalism after all.

References

Goldman, A. I. (2009), "Replies to Discussants", *Grazer philosophische Studien* 79: 245-288.

Goldman, A. I., and Olsson, E. J. (2009), "Reliabilism and the Value of Knowledge", in Haddock, A., Millar, A. and, Pritchard, D. H. (eds.), *Epistemic Value*, Oxford University Press.

Golinkoff, R., C. Mervis, and K. Hirsh-Pasek: (1994), "Early Object Labels: The Case for a Developmental Lexical Principles Framework", *Journal of Child Language* vol. 21: 125–156.

Jäger, C. "Reliabilism and the Value Problem", *Theoria*, this volume.

Kvanvig, J. L. (2003), *The Value of Knowledge and the Pursuit of Understanding*, Cambridge: Cambridge University Press.

Olsson, E. J. (2007), "Reliabilism, Stability, and the Value of Knowledge", *American Philosophical Quarterly*, vol. 44 (4): 343-355, 2007.

Olsson, E. J. (2008), "Knowledge, Truth, and Bullshit: Reflections on Frankfurt", *Midwest Studies in Philosophy* XXXII: 94-110.

Olsson, E. J. (2009), "In Defense of the Conditional Probability Solution to the Swamping Problem", *Grazer philosophische Studien* 79: 93-114.

Williamson, T. (2000), *Knowledge and Its Limits*, Oxford University Press.